UNITED STATES PATENT APPLICATION

FOR

LSF QUANTIZER FOR WIDEBAND SPEECH CODER

Inventor(s):   Dae-Sik Kim
                Song-In Choi
                Byung-Sik Yoon
                Hyung-Jung Kim
                Sang-Won Kang
                Sang-Hyun Chi

Prepared by:

LSF Quantizer for Wideband Speech Coder

BACKGROUND OF THE INVENTION

Field of the Invention

[0001] The present invention relates to a line spectral frequency (LSF) quantizer for a wideband speech coder. More specifically, the present invention relates to an LSF quantizer for a wideband speech coder that employs predictive pyramid vector quantization (PPVQ) and pyramid vector quantization (PVQ) usable for LSF quantization with a wideband speech quantizer.

Description of the Related Art

[0002] In general, it is of great importance to efficiently quantize an LSF coefficient indicating a correlation between short intervals of a speech signal for the sake of high-quality speech coding with a speech coder. The optimum linear predictive coefficient of a linear predictive coefficient (LPC) filter is calculated in a manner such that an input speech signal is divided by frames to minimize the energy of prediction errors by frame. The LPC filter of an AMR_WB (Adaptive Multi-Rate_Wideband) speech coder standardized as a wideband speech coder for a 3GPP IMT-2000 system by Nokia is a 16th-order all-pole filter that requires a certain number of bits to be allocated for quantization of the 16 linear predictive coefficients.

[0003] As an example, IS-96A QCELP (Qualcomm Code Excited Linear Prediction), a speech coding method for CDMA mobile communication systems, uses 25% of the total bits for LPC quantization, and an AMR_WB speech coder by Nokia uses 9.6 to 27.3 % of the total bits for the LPC quantization in nine modes. So far, many kinds of efficient LPC quantization methods have been developed and actually utilized in speech compressors. Direct quantization of the coefficients of the LPC filter is problematic in that the filter is too sensitive to the quantization error of the coefficients to guarantee stability of the LPC filter after coefficient

1

quantization. Accordingly, there is a need for converting the LPC to another parameter more suitable for quantization, such as a reflection coefficient or an LSF. In particular, the LSF value has a close relationship with the frequency characteristic of the speech signal so that most of the recent standard speech coders employ the LSF quantization method.

[0004] For efficient quantization, use is made of a correlation between frames of the LSF coefficient. Namely, the LSF of the current frame is not directly quantized but is predicted from that of the previous frame to quantize the prediction error. The LSF value is closely related to the frequency characteristic of the speech signal and thus is predictable in terms of time to obtain a considerably large prediction gain.

[0005] There are two prediction methods, one using an auto-regressive (AR) filter and the other using a moving average (MA) filter. The AR filter is superior in prediction performance but causes coefficient-transfer error propagation from one frame to another at a receiver. The MA filter is inferior in prediction performance to the AR filter but it is advantageous in that the effect of the transfer error is restrained over time. Accordingly, a prediction method with an MA filter is used in speech compressors such as AMR, CS-ACELP or EVRC that are utilized in environments in which many transfer errors occur, such as in radio communications.

[0006] The present invention solves the prediction error problem by use of both an AR predictor and a safety net. A quantization method using a correlation between neighboring LSF factors within a frame instead of LSF prediction between frames has also been developed. In particular, this method can promote the efficiency of quantization since the LSF values satisfy the order property.

[0007] It is impossible to quantize all vectors at the same time because of an extremely large vector table and a long retrieving time. To overcome this problem, a so-called split vector quantization (SVQ) method is suggested wherein the total vector is split into several subvectors, which are independently quantized. For example, the size of the vector table is 10×10^{20} in 10th-order vector quantization using 20 bits, but it is no more than 5

$\times 10^{20} \times 2$ in split vector quantization where the vector is split into two 5th-order subvectors to which 10 bits are independently allocated. Splitting the vector into more subvectors reduces the size of the vector table to save memory space, and hence the retrieving time, but it does not make the most of the correlation between vector values so it deteriorates performance.

[0008] With the vector split into ten 1st-order vectors, for example, the 10th-order vector quantization becomes scalar quantization. Assuming that split vector quantization is used to qauntize the LSF directly without LSF prediction between 20msec frames, 24 bits are required to attain the quantization performance. The split vector quantization method, in which the respective subvectors are independently quantized, causes a problem in that it cannot make the most of the correlation between the subvectors, hence it fails to optimize the total vector. Examples of other quantization methods recently developed include multi-stage vector quantization, a selective vector quantization method using two tables, and a linked split vector quantization method wherein a table to be used is selected with reference to the boundary values of the individual subvectors.

[0009] Although a general vector quantizer is required to store code books, the split vector quantizer has only to store the index of code books and enable ready calculation of the output vector without comparing the output vector with all other output codes possible in coding.

[0010] In general, the lattice is a set of nth-order vectors defined as Equation 1:

[0011] [Equation 1]

[0012] $\Lambda = \{x \mid x = c_1 a_1 + c_2 a_2 + ... + c_n a_n\}$

[0013] The split vector quantizer is largely classified into a uniform split vector quantizer and a pseudo-uniform split vector quantizer, and includes, depending on the type of code book, a spherical split vector quantizer or a pyramid split vector quantizer. The spherical split vector quantizer is suitable for a source having a Gaussian distribution, the pyramid split vector quantizer being suitable for a source having a Laplacian distribution.

3

SUMMARY OF THE INVENTION

[0014] It is an object of the present invention to provide an LSF quantizer for a wideband speech coder that reduces the size of memory and the computational complexity for retrieval of code books required in LPC quantization with an increase in the LPC order, and that decreases the number of outliers, with enhanced performance.

[0015] In one aspect of the present invention, an LSF (Line Spectral Frequency) quantizer for a wideband speech coder comprises: a subtracter for receiving an input LSF coefficient vector and removing a DC component from it; a memory-based vector quantizer and a memoryless vector quantizer for respectively receiving the DC component removed LSF coefficient vector and independently quantizing the same; a switch for receiving quantized vectors respectively quantized by the memory-based vector quantizer and the memoryless vector quantizer, selecting a quantized vector that has less quantized error that is a difference between the received quantized vector and the input LSF coeffcient vector from among the received quantized vectors, and outputting the same; and an adder for adding the quantized vector selected by the switch to the DC component of the LSF coefficient vector.

BRIEF DESCRIPTION OF THE DRAWINGS

[0016] The accompanying drawing, which is incorporated in and constitutes a part of the specification, illustrates an embodiment of the invention, and, together with the description, serves to explain the principles of the invention:

[0017] FIG. 1 is a schematic of an LSF quantizer for a wideband speech coder in accordance with an embodiment of the present invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

4

[0018] In the following detailed description, only the preferred embodiment of the invention has been shown and described, simply by way of illustration of the best mode contemplated by the inventor(s) of carrying out the invention. As will be realized, the invention is capable of modification in various obvious respects, all without departing from the invention. Accordingly, the drawing and description are to be regarded as illustrative in nature, and not restrictive.

[0019] Hereinafter, a detailed description will be given to an LSF quantizer for a wideband speech coder in accordance with an embodiment of the present invention with reference to the accompanying drawing.

[0020] For LSF quantization, an AMR_WB speech coder uses an S-MSVQ (Split-Multi Stage VQ) structure in which the DC component is removed, and a 16th-order prediction error vector, i.e., a difference value between a 16th-order LSF coefficient and a vector predicted by a primary MA predictor, is split into one 9th-order subvector and one 7th-order subvector for vector quantization, the 9th-order subvector being further split into three 3rd-order subvectors, and the 7th-order subvector being further split into one 3rd-order subvector and one 4th-order subvector. Such an S-MSVQ structure is to reduce the size of the memory and the code-book retrieving time required for 46-bit LSF coefficient quantization, and actually needs a relatively smaller memory and less computational complexity for retrieval of code books compared to the full VQ structure. But the S-MSVQ structure still requires a large memory ($2^8 + 2^8 + 2^6 + 2^7 + 2^7 + 2^5 + 2^5$) and a great deal of computational complexity because of complexity in retrieving code books.

[0021] For LSF quantization, the DC component is removed from the LSF value, and the LSF coefficient vector removed of the DC component is input to both a memory-based split quantizer (i.e., predictive PVQ) and a memoryless split quantizer (i.e., PVQ). The memory-based split quantizer (predictive PVQ), which is designed for fine quantization, pyramid-vector-quantizes an error vector that is a difference between a vector predicted by the primary AR predictor and an input vector. The memoryless split

5

quantizer, which is designed to reduce the number of outliers, directly pyramid-vector-quantizes the input vector. A candidate vector that minimizes an Euclidean distance from the original input vector from among two candidate vectors qunatized by the two qunatizers is selected to be a final quantized vector . Accordingly, the quantizer of the present invention has a strong point in that it provides the characteristics of both the memory-based split quantizer for fine quantization and the memoryless split quantizer for reducing the number of outliers.

[0022] The PVQ performance becomes favorable when the order of the input vector is high enough. That is, when the order of the input vector is more than about 20, the value $\|F(n)\|$ approximates a constant irrespective of the value of n. Otherwise, when the order of the input vector is below 20, the value $\|F(n)\|$ does not approximate a constant because of the large distribution of $\|F(n)\|$. This causes error propagation in quantization using a single pyramid. To solve this problem, there is suggested a product code PVQ (PCPVQ) that normalizes an input vector, quantizes it with a single pyramid and indexes the quantized pyramid using a normalized factor, $\hat{\gamma} = Q(\|F(n)\|)$. Here, $Q(\cdot)$ represents a scalar quantizer. When $\hat{c}(n) = PVQ(\hat{v}(n))$ is the output vector of PVQ and $\hat{\gamma} = Q(\|F(n)\|)$ is the output value of the scalar quantizer, the output vector of the product code PVQ, $\hat{c}_{PCPVQ}(n)$ is given by Equation 2:

[0023] [Equation 2]

[0024] $\hat{c}_{PCPVQ}(n) = \hat{\gamma} \leftarrow \hat{c}(n)$

[0025] This has an effect of using as many pyramids as quantization levels of the scalar quantizer. When the bit rate per average vector order of PVQ is R_p and the bit rate assigned to the scalar quantizer is R_γ, the total bit rate R satisfies Equation 3:

[0026] [Equation 3]

[0027] $R_p L + R_\gamma = RL$

[0028] FIG. 1 is a block diagram of a wideband LSF quantizer using a

6

memory-based predictive pyramid VQ and a memoryless pyramid VQ in accordance with an embodiment of the present invention.

[0029] The wideband LSF quantizer comprises: a subtracter 11 for receiving an input LSF coefficient vector and removing the DC component ; a memory-based PVQ 12 and a memoryless PVQ 13 for receiving the DC component-removed LSF coefficient vector R(n) and quantizing the same; a switch 14 for selecting the one of the vectors quantized by the memory-based PVQ 12 and the memoryless PVQ 13 that has the shorter Euclidean distance from the input LSF coefficient vector, and outputting the same; and an adder 15 for adding the vector selected by the switch 14 to the DC component of the LSF coefficient vector.

[0030] As described previously, the LSF coefficient quantizer for an AMR_WB speech coder using both a split VQ and a multi-stage VQ requires a relatively smaller memory and less computational complexity for retrieval of code books compared to the full VQ, but it still needs a large memory and a great deal of computational complexity. Additionally, the memory VQ structure causes error propagation. To solve this problem, the present invention uses a split vector quantizer that reduces the number of outliers and provides a simple coding procedure with a small memory. In particular, the present invention suggests a PVQ LSF coefficient quantizer using a pyramid split vector quantizer suitable for quantization of Laplacian signals, considering that the distribution of LSF coefficients has a characteristic of Laplacian signals.

[0031] An operation of the quantizer shown in FIG. 1 is as follows. Upon receiving an LSF coefficient vector, the subtracter 11 removes the DC component from the LSF coefficient vector. The DC component-removed LSF coefficient vector is fed into both the memory-based PVQ 12 and the memoryless PVQ 13 to be independently quantized. The memory-base PVQ, i.e., the predictive pyramid VQ, predicts the input vector using a primary AR predictor, and uses the pyramid VQ (PVQ) to quantize a prediction error vector which is a difference between the predicted vector and the input vector. The memoryless PVQ, i.e., pyramid VQ (PVQ),

7

quantizes the input vector in the full vector format using a pyramid VQ designed for focusing on the outliers. The quantized error, that is, a difference between each of the quantized vectors and the input vector, is determined in terms of Euclidean distance, so that a candidate vector having a less quantized error is selected as the quantized vector. The quantized values obtained by the two quantizers in a quantization program produce two Euclidean distances as error values between the value before quantization and the quantized value. The quantizer of the present invention selects the one of the two quantized values that has the shorter Euclidean distance.

[0032] As described above, the present invention employs a split vector quantizer of a novel structure as an LSF coefficient quantizer for an AMR_WB speech coder in order to reduce the size of memory and computational complexity for retrieval of code books, and to improve the bit rate and the spectral distortion (SD).

[0033] While this invention has been described in connection with what is presently considered to be the most practical and preferred embodiment, it is to be understood that the invention is not limited to the disclosed embodiments, but, on the contrary, is intended to cover various modifications and equivalent arrangements included within the spirit and scope of the appended claims.

[0034] According to the present invention, as described above, the use of a split vector quantizer and a safety net in the LSF coefficient quantizer greatly reduces the size of the memory and the computational complexity for retrieval of code books without a deterioration of the SD performance. An experiment reveals that the total number of bits used to attain an SD performance of 1 dB using the above quantizer is no more than 39 bits, which is less by 7 bits than the 46 bits required by an AMR-WB speech coder.